Continuous Hands-off Control by CLOT Norm Minimization

Niharika Challapalli * **Masaaki Nagahara** ** **Mathukumalli Vidyasagar** ***

* *Department of Electrical Engineering University of Texas at Dallas Richardson, TX 75080, USA (e-mail: niharika15c@gmail.com)*
** *Institute of Environmental Science and Technology, The University of Kitakyushu, Hibikino 1-1, Wakamatsu-ku, Kitakyushu, Fukuoka 808-0135, JAPAN (e-mail: nagahara@ieee.org).*
*** *Department of Systems Engineering, University of Texas at Dallas, Richardson, TX 75080, USA, and Department of Electrical Engineering, Indian Institute of Technology Hyderabad, Kandi, Telangana, India 502285 (e-mail: m.vidyasagar@utdallas.edu, m.vidyasagar@iith.ac.in)*

Abstract: In this paper, we consider hands-off control via minimization of the CLOT (Combined L-One and Two) norm. The maximum hands-off control is the L^0-optimal (or the sparsest) control among all feasible controls that are bounded by a specified value and transfer the state from a given initial state to the origin within a fixed time duration. In general, the maximum hands-off control is a bang-off-bang control taking values of ± 1 and 0. For many real applications, such discontinuity in the control is not desirable. To obtain a continuous but still relatively sparse control, we propose to use the CLOT norm, a convex combination of L^1 and L^2 norms. We show by numerical simulation that the CLOT control is continuous and much sparser (i.e. has longer time duration on which the control takes 0) than the conventional EN (elastic net) control, which is a convex combination of L^1 and squared L^2 norms.

Keywords: Optimal control, convex optimization, sparsity, maximum hands-off control, bang-off-bang control

1. INTRODUCTION

Sparsity has recently emerged as an important topic in signal/image processing, machine learning, statistics, etc. If $y \in \mathbb{R}^m$ and $A \in \mathbb{R}^{m \times n}$ are specified with $m < n$, then the equation $y = Ax$ is underdetermined and has infinitely many solutions for x if A has rank m. Finding the sparsest solution (that is, the solution with the fewest number of nonzero elements) can be formulated as

$$\min_z \|z\|_0 \quad \text{subject to} \quad Az = b.$$

However, this problem is NP hard, as shown in (Natarajan, 1995). Therefore other approaches have been proposed for this purpose. This area of research is known as "sparse regression." One of the most popular is LASSO (Tibshirani, 1996), also referred to as basis pursuit (Chen et al., 1999), in which the ℓ^0-norm is replaced by the ℓ^1-norm. Thus the problem becomes

$$\min_z \|z\|_1 \quad \text{subject to} \quad Az = b.$$

The advantage of LASSO is that it is a convex optimization problem and therefore very large problems can be solved efficiently, for example by using the `Matlab`-based package `cvx` (Grant and Boyd, 2014). Moreover, under mild technical assumptions, the LASSO-optimal solution has no more than m nonzero components (Osborne et al., 2000). However, the exact location of the nonzero components is very sensitive to the vector y. To overcome this

deficiency, another approach known as the Elastic Net was proposed in (Zou and Hastie, 2005), where the ℓ^1 norm in LASSO is replaced by a weighted sum of ℓ^1 and squared ℓ^2 norms. This leads to the optimization problem

$$\min_z \lambda_1 \|z\|_1 + \lambda_2 \|z\|_2^2 \quad \text{subject to} \quad Az = b,$$

where λ_1 and λ_2 are positive weights such that $\lambda_1 + \lambda_2 = 1$. It is shown in (Zou and Hastie, 2005, Theorem 1) that the EN formulation gives the *grouping effect*; If two columns of the matrix A are highly correlated, then the corresponding components of the solution for x have nearly equal values. This ensures that the solution for x is not overly sensitive to small changes in y. The name "elastic net" is meant to suggest a stretchable fishing net that retains *all the big fish*.

During the past decade and a half, another research area known as "compressed sensing" has witnessed a great deal of interest. In compressed sensing, the matrix A is not specified; rather, the user gets to choose the integer m (known as the number of measurements), as well as the matrix A. The objective is to choose the matrix A as well as a corresponding "decoder" map $\Delta : \mathbb{R}^m \to \mathbb{R}^n$ such that, the unknown vector x is sparse and the measurement vector y equals Ax, then $\Delta(Ax) = x$ for all sufficiently sparse vectors x. More generally, if measurement vector $y = Ax + \eta$ where η is the measurement noise, and the vector x is nearly sparse (but not exactly sparse), then

the recovered vector $\Delta(Ax + \eta)$ should be sufficiently close to the true but unknown vector x. This is referred to as "robust sparse recovery." Minimizing the ℓ_1-norm is among the more popular decoders. See the books by (Elad, 2010), (Eldar and Kutyniok, 2012), and (Foucart and Rauhut, 2013) for the theory and some applications. Due to its similarity to the LASSO formulation of (Tibshirani, 1996), this approach to compressed sensing is also referred to as LASSO.

Until recently the situation was that LASSO achieves robust sparse recovery in compressed sensing, but did not achieve the grouping effect in sparse regression. On the flip side, EN achieves the grouping effect, but it was not known whether it achieves robust sparse recovery. A recent paper (Ahsen et al., 2016) sheds some light on this problem. It is shown in (Ahsen et al., 2016) that EN *does not achieve* robust sparse recovery. To achieve both the grouping effect in sparse regression as well as robust sparse recovery in compressed sensing, (Ahsen et al., 2016) has proposed the CLOT (Combined L-One and Two) formulation:

$$\min_z \lambda_1 \|z\|_1 + \lambda_2 \|z\|_2 \ \ \text{subject to} \ \ Az = b \ \text{and} \ \lambda_1 + \lambda_2 = 1.$$

The difference between EN and CLOT is the ℓ^2 norm term; EN has the squared ℓ^2 norm while CLOT has the pure ℓ^2 norm. This slight change leads to both the grouping effect and robust sparse recovery, as shown in (Ahsen et al., 2016).

In parallel with these advances in sparse regression and recovery of unknown sparse vectors, sparsity techniques have also been applied to control. Sparsity-promoting optimization has been applied to networked control in (Nagahara et al., 2014), where quantization errors and data rate can be reduced at the same time by sparse representation of control packets. Other examples of control applications include optimal controller placement by (Casas et al., 2012; Clason and Kunisch, 2012; Fardad et al., 2011), design of feedback gains by (Lin et al., 2013; Polyak et al., 2013), state estimation by (Charles et al., 2011), to name a few.

More recently, a novel control called the *maximum hands-off control* has been proposed in (Nagahara et al., 2016) for *continuous-time* systems. The maximum hands-off control is the L^0-optimal control (the control that has the minimum support length) among all feasible controls that are bounded by a fixed value and transfer the state from a given initial state to the origin within a fixed time duration. Such a control is effective for reduction of electricity or fuel consumption; an electric/hybrid vehicle shuts off the internal combustion engine (i.e. hands-off control) when the vehicle is stopped or the speed is lower than a preset threshold; see (Chan, 2007) for example. Railway vehicles also utilize hands-off control, often called *coasting control*, to cut electricity consumption; see (Liu and Golovitcher, 2003) for details. In (Nagahara et al., 2016), the authors have proved the theoretical relation between the maximum hands-off control and the L^1 optimal control under the assumption of normality. Also, important properties of the maximum hands-off control have been proved in (Ikeda and Nagahara, 2016) for the convexity of the value function, and in (Chatterjee et al., 2016) for the existence and the discreteness.

In general, the maximum hands-off control is a bang-off-bang control taking values of ± 1 and 0. For many real applications, such a discontinuity property is not desirable. To obtain a continuous but still sparse control, (Nagahara et al., 2016) has proposed to use a combined L^1 and *squared* L^2 minimization, like EN mentioned above. Let us call this control an EN control. As in the case of EN in the vector optimization, the EN control often shows much less sparse (i.e. has a larger L^0 norm) than the maximum hands-off control. Then, we proposed to use the CLOT norm, a convex combination of L^1 and *non-squared* L^2 norms. We call the minimum CLOT-norm control the CLOT control. We show by numerical simulation that the CLOT control is continuous and much sparser (i.e. has longer time duration on which the control takes 0) than the conventional EN control.

The remainder of this article is organized as follows. In Section 2, we formulate the control problem considered in this paper. In Section 3, we give a discretization method to numerically compute the optimal control. Results of the numerical computations on a variety of problems are presented in Section 4. These examples illustrate the advantages of the CLOT control compared with the maximum hands-off control and the EN control. We present some conclusions in Section 5.

Notation

Let $T > 0$ and $m \in \mathbb{N}$. For a continuous-time signal $u(t) \in \mathbb{R}$ over a time interval $[0, T]$, we define its L^p $(p \geq 1)$ and L^∞ norms respectively by

$$\|u\|_p \triangleq \left\{ \int_0^T |u(t)|^p dt \right\}^{1/p}, \ \|u\|_\infty \triangleq \sup_{t \in [0,T]} |u(t)|.$$

We denote the set of all signals with $\|u\|_p < \infty$ by $L^p[0, T]$ for $p \geq 1$ or $p = \infty$. We define the L^0 norm of a signal $u(t)$ on the interval $[0, T]$ as

$$\|u\|_0 \triangleq \int_0^T \phi_0(u(t)) dt,$$

where ϕ_0 is the L^0 kernel function defined by

$$\phi_0(\alpha) \triangleq \begin{cases} 1, & \text{if } \alpha \neq 0, \\ 0, & \text{if } \alpha = 0 \end{cases} \quad (1)$$

for a scalar $\alpha \in \mathbb{R}$. The L^0 norm can be represented by

$$\|u\|_0 = \mu_{\mathrm{L}}\big(\mathrm{supp}(u)\big),$$

where $\mathrm{supp}(u)$ is the support of the signal u, and μ_{L} is the Lebesgue measure on $\mathbb{R}$.

2. PROBLEM FORMULATION

Let us consider a linear time-invariant system described by

$$\frac{dx}{dt}(t) = Ax(t) + Bu(t), \ t \geq 0, \ x(0) = \xi. \quad (2)$$

Here we assume that $x(t) \in \mathbb{R}^n$, $u(t) \in \mathbb{R}$, and the initial state $x(0) = \xi$ is fixed and given. The control objective is to drive the state $x(t)$ from $x(0) = \xi$ to the origin at time $T > 0$, that is

$$x(T) = 0. \quad (3)$$

We limit the control $u(t)$ to satisfy

$$\|u\|_\infty \leq U_{\max} \quad (4)$$

for fixed $U_{\max} > 0$.

If the system (2) is controllable and the final time T is larger than the optimal time T^* (the minimal time in which there exist a control $u(t)$ that drives $x(t)$ from $x(0) = \xi$ to the origin; see (Hermes and Lasalle, 1969)), then there exists at least one $u(t) \in L^\infty[0,T]$ that satisfies equations (2), (3), and (4). Let us call such a control a *feasible* control. From (2) and (3), any feasible control $u(t)$ on $[0,T]$ satisfies

$$0 = x(T) = e^{AT}\xi + \int_0^T e^{A(T-t)}Bu(t)dt,$$

or

$$\int_0^T e^{-At}Bu(t)dt + \xi = 0. \tag{5}$$

Define a linear operator $\Phi : L^\infty[0,T] \to \mathbb{R}^n$ by

$$\Phi u \triangleq \int_0^T e^{-At}Bu(t)dt, \quad u \in L^\infty[0,T].$$

By this, we define the set $\mathcal{U}$ of the feasible controls by

$$\mathcal{U} \triangleq \{u \in L^\infty : \Phi u + \xi = 0, \|u\|_\infty \le 1\}. \tag{6}$$

The problem of the maximum hands-off control is then described by

$$\min_u \|u\|_0 \quad \text{subject to} \quad u \in \mathcal{U}. \tag{7}$$

The L^0 problem (7) is very hard to solve since the L^0 cost function is non-convex and discontinuous. For this problem, (Nagahara et al., 2016) has shown that the L^0 optimal control in (7) is equivalent to the following L^1 optimal control:

$$\min_u \|u\|_1 \quad \text{subject to} \quad u \in \mathcal{U}, \tag{8}$$

if the plant is normal, that is, if the (2) is controllable and the matrix A is nonsingular. Let us call the L^1 optimal control as the *LASSO control*. If the plant is normal, then the LASSO control is in general a *bang-off-bang* control that is piecewise constant taking values in $\{0, \pm 1\}$. The discontinuity of the LASSO solution is not desirable in real applications, and a smoothed solution is also proposed in (Nagahara et al., 2016) as

$$\min_u \|u\|_1 + \lambda\|u\|_2^2 \quad \text{subject to} \quad u \in \mathcal{U}, \tag{9}$$

where $\lambda > 0$ is a design parameter for smoothness. Let us call this control the *EN (elastic net) control*. In (Nagahara et al., 2016), it is proved that the solution of (9) is a continuous function on $[0,T]$.

While the EN control is continuous, it is shown by numerical experiments that the EN control is not sometimes sparse. This is an analogy of the EN for finite-dimensional vectors that EN does not achieve robust sparse recovery. Borrowing the idea of CLOT in (Ahsen et al., 2016), we define the CLOT optimal control problem by

$$\min_u \|u\|_1 + \lambda\|u\|_2 \quad \text{subject to} \quad u \in \mathcal{U}. \tag{10}$$

We call this optimal control the *CLOT control*.

3. DISCRETIZATION

Since the problems (8)–(10) are infinite dimensional, we should approximate it to finite dimensional problems. For this, we adopt the time discretization.

First, we divide the time interval $[0,T]$ into N subintervals, $[0,T] = [0,h) \cup \cdots \cup [(N-1)h, Nh]$, where h is the discretization step (or the sampling period) such that $T = Nh$. We assume that the state $x(t)$ and the control $u(t)$ in (2) are constant over each subinterval. On the discretization grid, $t = 0, h, \ldots, Nh$, the continuous-time system (2) is described as

$$x_d[m+1] = A_d x_d[m] + B_d u_d[m], \quad m = 0, 1, \ldots, N-1, \tag{11}$$

where $x_d[m] \triangleq x(mh)$, $u_d[m] \triangleq u(mh)$, and

$$A_d \triangleq e^{Ah}, \quad B_d \triangleq \int_0^h e^{At}Bdt. \tag{12}$$

Define the control vector

$$\boldsymbol{u}_d \triangleq [u_d[0], u_d[1], \ldots, u_d[N-1]]^\top. \tag{13}$$

Note that the final state $x(T)$ can be described as

$$x(T) = x_d[N] = A_d^N \xi + \Phi_N \boldsymbol{u}_d, \tag{14}$$

where

$$\Phi_N \triangleq \left[A_d^{N-1}B_d, \ A_d^{N-2}B_d, \ \ldots, \ B_d\right]. \tag{15}$$

Then the set $\mathcal{U}$ in (6) is approximately represented by

$$\mathcal{U}_N \triangleq \left\{\boldsymbol{u}_d \in \mathbb{R}^N : A_d^N \xi + \Phi_N \boldsymbol{u}_d = 0, \ \|\boldsymbol{u}_d\|_\infty \le 1\right\}. \tag{16}$$

Next, we approximate the L^1 norm of u by

$$\begin{aligned}
\|u\|_1 &= \int_0^T |u(t)|dt \\
&= \sum_{m=0}^{N-1} \int_{mh}^{(m+1)h} |u(t)|dt \\
&\approx \sum_{m=0}^{N-1} \int_{mh}^{(m+1)h} |u_d[m]|dt \\
&= \sum_{m=0}^{N-1} |u_d[m]|h \\
&= \|\boldsymbol{u}_d\|_1 h.
\end{aligned} \tag{17}$$

In the same way, we obtain approximation of the L^2 norm of u as

$$\|u\|_2^2 = \int_0^T |u(t)|^2 dt \approx \|\boldsymbol{u}_d\|_2^2 h. \tag{18}$$

Finally, the optimal control problems (8), (9) and (10) can be approximated by

$$\min_{\boldsymbol{u}_d \in \mathbb{R}^N} h\|u\|_1 \quad \text{subject to} \quad \boldsymbol{u}_d \in \mathcal{U}_N \tag{19}$$

$$\min_{\boldsymbol{u}_d \in \mathbb{R}^N} h\|\boldsymbol{u}_d\|_1 + h\lambda\|\boldsymbol{u}_d\|_2^2 \quad \text{subject to} \quad \boldsymbol{u}_d \in \mathcal{U}_N \tag{20}$$

$$\min_{\boldsymbol{u}_d \in \mathbb{R}^N} h\|\boldsymbol{u}_d\|_1 + \sqrt{h}\lambda\|\boldsymbol{u}_d\|_2 \quad \text{subject to} \quad \boldsymbol{u}_d \in \mathcal{U}_N \tag{21}$$

The optimization problems are convex and can be efficiently solved by numerical software packages such as `cvx` with Matlab; see (Grant and Boyd, 2014) for details.

4. NUMERICAL EXAMPLES

In this section we present numerical results from applying the CLOT norm minimization approach to seven different plants, and compare the results with those from applying LASSO and EN.

No.	Plant	Poles	T	$x(0)$	λ	Figs.
1	$P_1(s)$	$(0,0,0,0)$	20	$\mathbf{e}_4$	1	1 , 2
2	$P_1(s)$	$(0,0,0,0)$	20	$\mathbf{e}_4$	0.1	3 , 4
3	$P_2(s)$	$-0.025 \pm j$	20	$\mathbf{e}_2$	0.1	5 , 6
4	$P_2(s)$	$-0.025 \pm j$	20	$(10,1)^\top$	0.1	7 , 8
5	$P_3(s)$	$-1 \pm 0.2j, \pm j$	20	$\mathbf{e}_4$	0.1	9, 10
6	$P_4(s)$	$-1 \pm 0.2j,$ $-0.3 \pm j$	20	$\mathbf{e}_4$	0.1	11, 12
7	$P_5(s)$	$-5 \pm j,$ $-0.3 \pm 2j$, $-1 \pm 2\sqrt{2}j$	20	$\mathbf{e}_6$	0.1	13, 14
8	$P_6(s)$	$0,0,0,0,\pm j$	40	$\mathbf{e}_6$	0.1	15, 16
9	$P_7(s)$	$0,0,0,0,\pm j$	40	$\mathbf{e}_6$	0.1	17, 18

Table 1. Details of various plants studied.

4.1 Details of Various Plants Studied

For the reader's convenience, the details of the various plants are given in Table 1. The figure numbers show where the corresponding computational results can be found. Some conventions are adopted to reduce the clutter in the table, as described next. All plants are of the form

$$P(s) = \frac{n(s)}{d(s)}, n(s) = \prod_{i=1}^{n_z}(s - z_i), d(s) = \prod_{i=1}^{n_p}(s - p_i).$$

To save space in the table, the plant zeros are not shown; $P_3(s)$ has a zero at $s = -2$, $P_6(s)$ has a zero at $s = 2$, while $P_7(s)$ has zeros at $s = 1, 2$. The remaining plants do not have any zeros, so that the plant numerator equals one.

Once the plant zeros and poles are specified, the plant numerator and denominator polynomials n, d were computed using the Matlab command `poly`. Then the transfer function was computed as `P = tf(n,d)`, and the state space realization was computed as `[A,B,C,D] = ssdata(P)`. The maximum control amplitude is taken 1, so that the control must satisfy $|u(t)| \leq 1$ for $t \in [0,T]$. To save space, we use the notation $\mathbf{e}_l$ to denote an l-column vector whose elements all equal one. Note that in all but one case, the initial condition equals $\mathbf{e}_n$ where n is the order of the plant.

Note that, with $T = 20$, the problems with plants $P_6(s)$ and $P_7(s)$ are not feasible (meaning that T is smaller than the minimum time needed to reach the origin); this is why we took $T = 40$.

All optimization problems were solved after discretizing the interval $[0, T]$ into both 2,000 as well as 4,000 samples, to examine whether the sampling time affects the sparsity density of the computed optimal control.

4.2 Plots of Optimal State and Control Trajectories

The plots of the ℓ^2-norm (or Euclidean norm) of the state vector trajectory and the control signal for all these examples are shown in the next several plots.

We begin with the plant $P_1(s)$, the fourth-order integrator. Figures 1 and 2 show the state and control trajectories when $\lambda = 1$. The same system is analyzed using a smaller value of $\lambda = 0.1$. One would expect that the resulting control signals would be more sparse with a smaller λ, and this is indeed the case. The results are shown in Figures 3 and 4. Based on the observation that the control signal
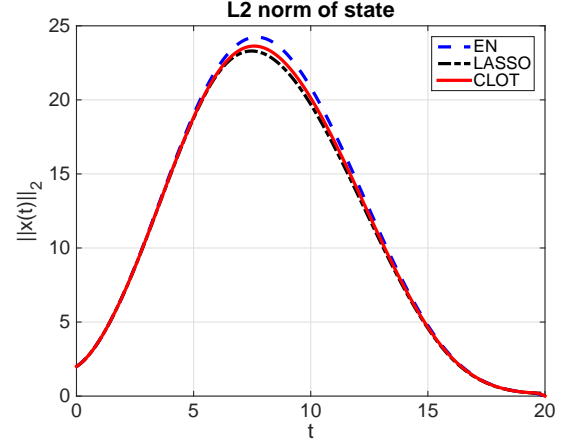

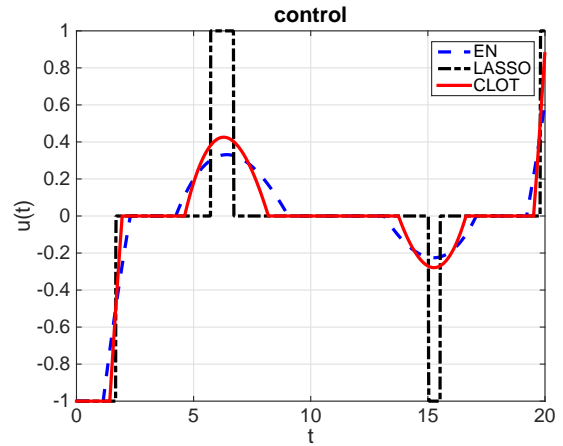Fig. 1. State trajectory for the plant $P_1(s)$ with the initial state $(1,1,1,1)^\top$ and $\lambda = 1$.

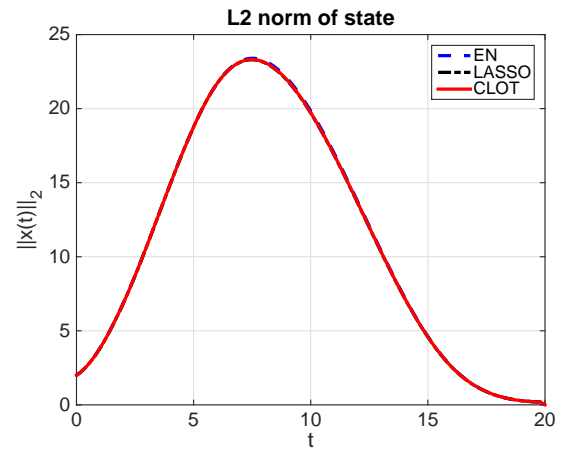Fig. 2. Control trajectory for the plant $P_1(s)$ with the initial state $(1,1,1,1)^\top$ and $\lambda = 1$.

Fig. 3. State trajectory for the plant $P_1(s)$ with the initial state $(1,1,1,1)^\top$ and $\lambda = 0.1$.

becomes more sparse with $\lambda = 0.1$ than with $\lambda = 1$, all the other plants are analyzed with $\lambda = 0.1$.

Figures 5 and 6 display the state trajectory and the control trajectories of the plant $P_2(s)$ (damped harmonic oscillator) when the initial state is $(1,1)^\top$. Figures 7 and
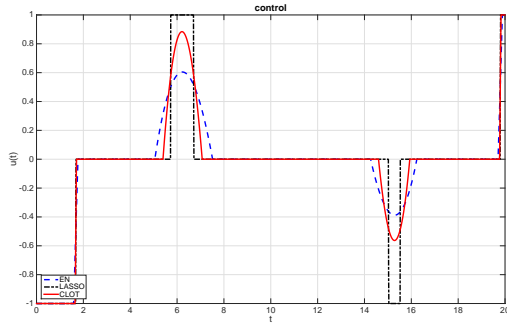
Fig. 4. Control trajectory for the plant $P_1(s)$ with the initial state $(1, 1, 1, 1)^\top$ and $\lambda = 0.1$.
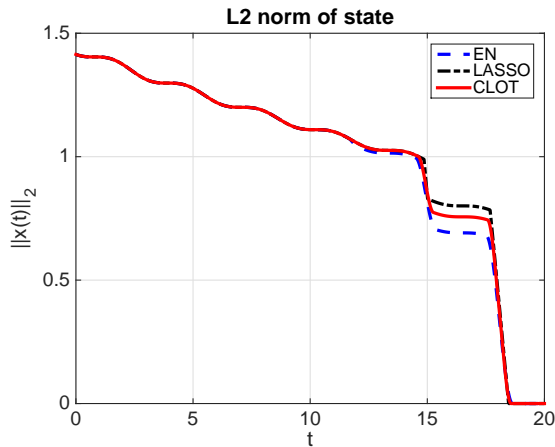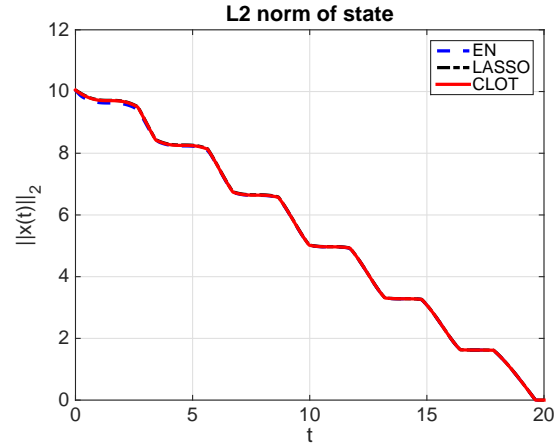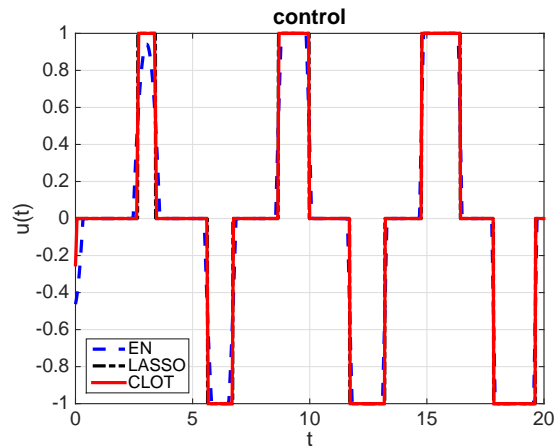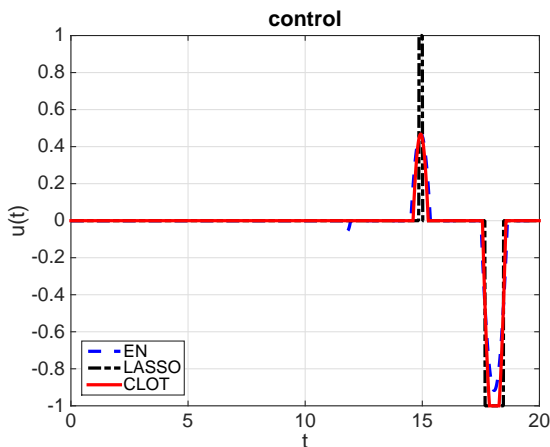

Fig. 5. State trajectory for the the plant $P_2(s)$ with the initial state $(1, 1)^\top$ and $\lambda = 0.1$.

Fig. 6. Control trajectory for the the plant $P_2(s)$ with the initial state $(1, 1)^\top$ and $\lambda = 0.1$.
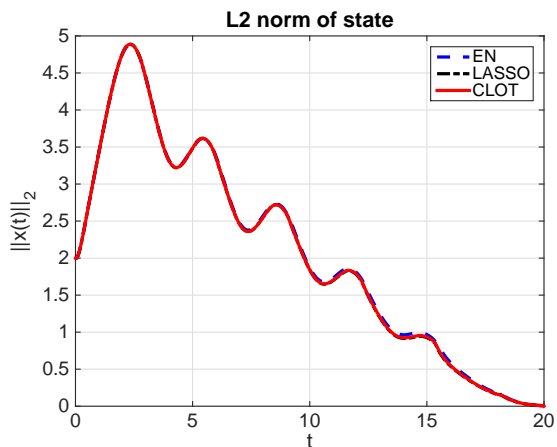
8 show the state and control trajectories with the initial state $(10, 1)^\top$. It can be seen that, with this intial state, the control signal changes sign more frequently.

To compare the sparsity densities of the three control signals, we compute the fraction of time that each signal is nonzero. In this connection, it should be noted that the LASSO control signal is the solution of a *linear programming* problem; consequently its components *exactly* equal zero at many time instants. In contrast, the EN

Fig. 7. State trajectory for the plant $P_2(s)$ with the initial state $(10, 1)^\top$ and $\lambda = 0.1$.

Fig. 8. Control trajectory for the plant $P_2(s)$ with the initial state $(10, 1)^\top$ and $\lambda = 0.1$.

Fig. 9. State trajectory for the plant $P_3(s)$ with the initial state $(1, 1, 1, 1)^\top$ and $\lambda = 0.1$.

and CLOT control signals are the solutions of *convex* optimization problems. Consequently, there are many time instants when the control signal is "small" without being smaller than the machine zero. Therefore, to compute the sparsity density, we applied a threshold of 10^{-4}, and treated a component of a control signal as being zero if
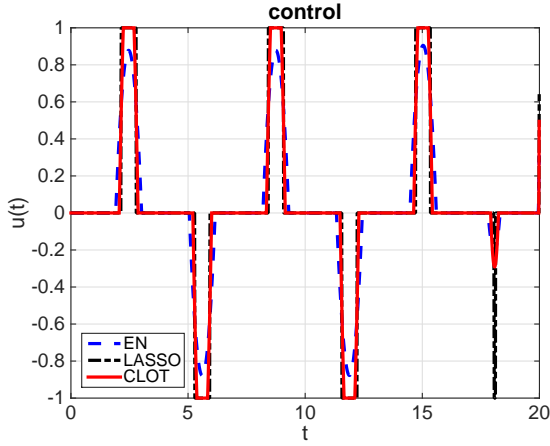
Fig. 10. Control trajectory for the plant $P_3(s)$ with the initial state $(1, 1, 1, 1)^\top$ and $\lambda = 0.1$.
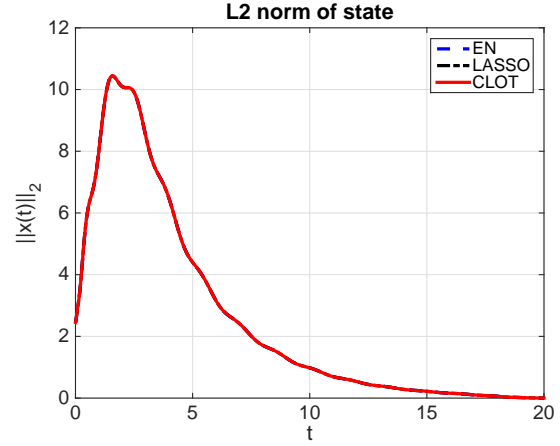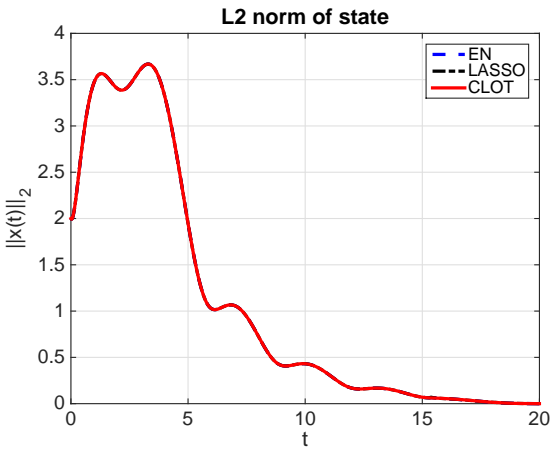

Fig. 11. State trajectory for the plant $P_4(s)$ with the initial state $(1, 1, 1, 1)^\top$ and $\lambda = 0.1$.
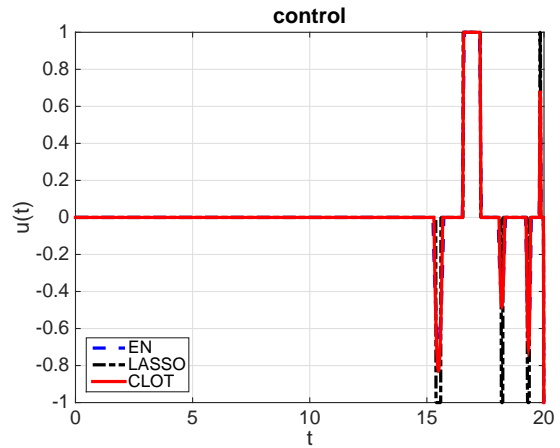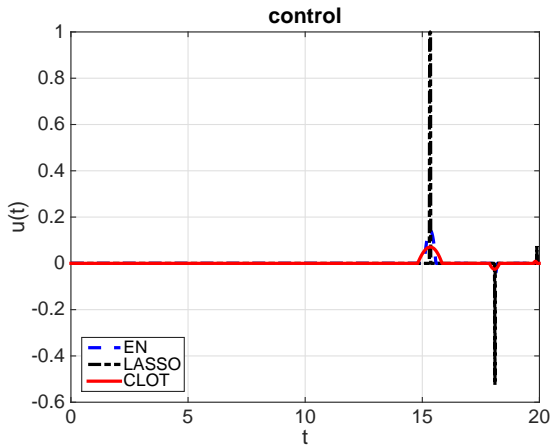

Fig. 12. Control trajectory for the plant $P_4(s)$ with the initial state $(1, 1, 1, 1)^\top$ and $\lambda = 0.1$.

Fig. 13. State trajectory for the plant $P_5(s)$ with the initial state $(1, 1, 1, 1, 1, 1)^\top$ and $\lambda = 0.1$.

Fig. 14. Control trajectory for the plant $P_5(s)$ with the initial state $(1, 1, 1, 1, 1, 1)^\top$ and $\lambda = 0.1$.
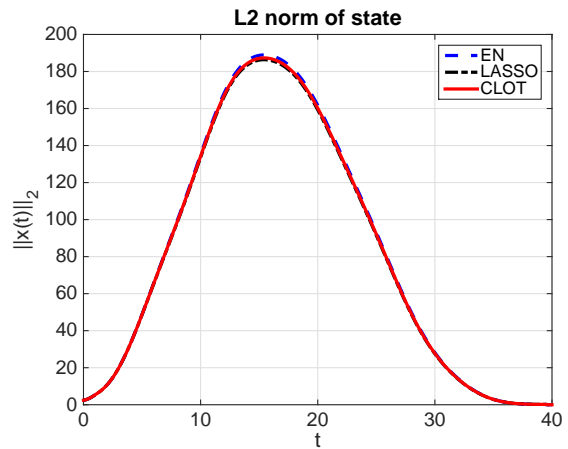

Fig. 15. State trajectory for the plant $P_6(s)$ with the initial state $(1, 1, 1, 1, 1, 1)^\top$ and $\lambda = 0.1$.

its magnitude is smaller than this threshold. With this convention, the sparsity densities of the various control signals are as shown in Table 2. From this table it can be seen that the control signal generated using CLOT norm minimization has significantly lower sparsity density compared to that of EN, and is not much higher than

that of LASSO. Also, as expected, the sparsity density of LASSO does not change with λ, whereas the sparsity densities of both EN and CLOT decrease as λ is decreased. For this reason, in other examples we present only the results for $\lambda = 0.1$.
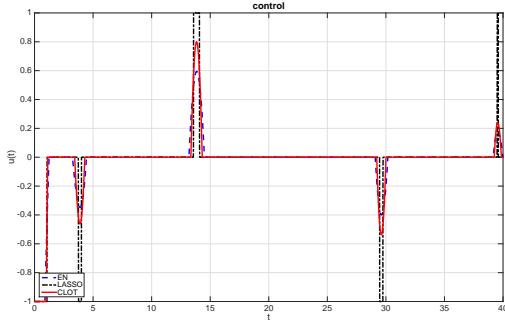
Fig. 16. Control trajectory for the plant $P_6(s)$ with the initial state $(1,1,1,1,1,1)^\top$ and $\lambda = 0.1$.
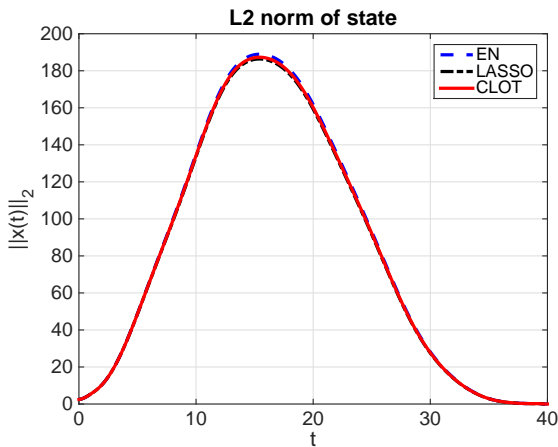

Fig. 17. State trajectory for the plant $P_7(s)$ with the initial state $(1,1,1,1,1,1)^\top$ and $\lambda = 0.1$.
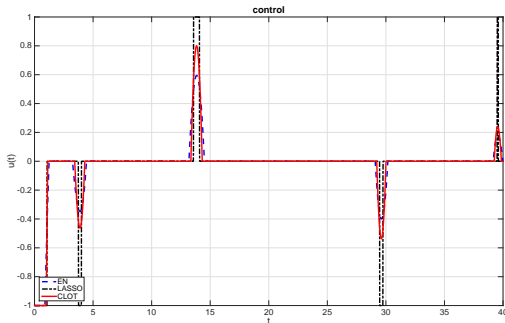

Fig. 18. Control trajectory for the plant $P_7(s)$ with the initial state $(1,1,1,1,1,1)^\top$ and $\lambda = 0.1$.

λ	LASSO	EN	CLOT
$\lambda = 1$	0.1725	0.6050	0.5900
$\lambda = 0.1$	0.1725	0.3795	0.2665

Table 2. Sparsity indices of the control signals from various algorithms for the plant $P_1(s)$ (fourth-order integrator) with the initial state $(1,1,1,1)$.

4.3 Comparison of Sparsity Densities

In this subsection we analyze the sparsity densities, that is, the fraction of samples that are nonzero, using the

No.	LASSO	EN	CLOT
1	0.1690	0.5915	0.4475
2	0.1690	0.3270	0.2480
3	0.0480	0.1155	0.0830
4	0.4055	0.5555	0.4225
5	0.1655	0.3050	0.2180
6	0.0040	0.0395	0.0805
7	0.0595	0.1100	0.0845
8	0.0568	0.1438	0.1125
9	0.0568	0.1438	0.1125

Table 3. Sparsity densities for optimal controllers produced by various methods

three methods LASSO, EN, and CLOT. The advantage of using the sparsity *density* instead of the sparsity *count* (the absolute number of nonzero entries) is that when the sample time is reduced, the sparsity count would increase, whereas we would expect the sparsity density to remain the same. As explained above, we have applied a threshold of 10^{-4} in computing the sparsity densities of various control signals.

Table 3 shows the sparsity densities for the nine examples studied in Table 1, in the same order. From this table it can be seen that the CLOT norm-based control signal is always more sparse than the EN-based control signal. Indeed, in some cases the sparsity density of the CLOT control is comparable to that of the LASSO control.

We also increased the number of samples from 2,000 to 4,000, and the optimal values changed only in the third significant figure in almost all examples for all three methods. Therefore the figures in Table 3 are essentially equal to the Lebesgue measure of the support set divided by T.

5. CONCLUSIONS

In this article, we propose the CLOT norm-based control that minimizes the weighted sum of L^1 and L^2 norms among feasible controls, to obtain a continuous control signal that is sparser than the EN control introduced in (Nagahara et al., 2016). We have shown a discretization method, by which the CLOT optimal control problem can be solved via finite-dimensional convex optimization. Numerical experiments have shown the advantage of the CLOT control compared with the LASSO and EN controls. Future work includes the analysis of the sparsity and continuity of the CLOT control as a continuous-time signal.

REFERENCES

Ahsen, M.E., Challapalli, N., and Vidyasagar, M. (2016). Two new approaches to compressed sensing exhibiting

both robust sparse recovery and the gourping effect. *Journal of Machine Learning Research.* (submitted; preprint at arXiv 1410:8229).

Casas, E., Clason, C., and Kunisch, K. (2012). Approximation of elliptic control problems in measure spaces with sparse solutions. *SIAM J. Control Optim.*, 50, 1735–1752.

Chan, C. (2007). The state of the art of electric, hybrid, and fuel cell vehicles. *Proc. IEEE*, 95(4), 704–718.

Charles, A., Asif, M., Romberg, J., and Rozell, C. (2011). Sparsity penalties in dynamical system estimation. In *45th Annual Conference on Information Sciences and Systems (CISS)*, 1–6.

Chatterjee, D., Nagahara, M., Quevedo, D.E., and Rao, K.M. (2016). Characterization of maximum hands-off control. *Systems & Control Letters*, 94, 31–36.

Chen, S.S., Donoho, D.L., and Saunders, M.A. (1999). Atomic decomposition by basis pursuit. *SIAM J. Sci. Comput.*, 20(1), 33–61.

Clason, C. and Kunisch, K. (2012). A measure space approach to optimal source placement. *Comput. Optim. Appl,*, 53, 155–171.

Elad, M. (2010). *Sparse and Redundant Representations.* Springer.

Eldar, Y.C. and Kutyniok, G. (2012). *Compressed Sensing: Theory and Applications.* Cambridge University Press.

Fardad, M., Lin, F., and Jovanović, M. (2011). Sparsity-promoting optimal control for a class of distributed systems. In *American Control Conference (ACC), 2011*, 2050–2055.

Foucart, S. and Rauhut, H. (2013). *A Mathematical Introduction to Compressive Sensing.* Birkhäuser.

Grant, M. and Boyd, S. (2014). CVX: Matlab software for disciplined convex programming, version 2.1. http://cvxr.com/cvx.

Hermes, H. and Lasalle, J.P. (1969). *Functional Analysis and Time Optimal Control.* Academic Press.

Ikeda, T. and Nagahara, M. (2016). Value function in maximum hands-off control for linear systems. *Automatica*, 64, 190–195.

Lin, F., Fardad, M., and Jovanovic, M.R. (2013). Design of optimal sparse feedback gains via the alternating direction method of multipliers. *IEEE Trans. Autom. Control*, 58(9), 2426–2431.

Liu, R. and Golovitcher, I.M. (2003). Energy-efficient operation of rail vehicles. *Transportation Research Part A: Policy and Practice*, 37(10), 917–932.

Nagahara, M., Quevedo, D.E., and Nevsić, D. (2016). Maximum hands-off control: a paradigm of control effort minimization. *IEEE Trans. Autom. Control*, 61(3), 735–747.

Nagahara, M., Quevedo, D., and Østergaard, J. (2014). Sparse packetized predictive control for networked control over erasure channels. *IEEE Trans. Autom. Control*, 59(7), 1899–1905.

Natarajan, B.K. (1995). Sparse approximate solutions to linear systems. *SIAM J. Comput.*, 24(2), 227–234.

Osborne, M.R., Presnell, B., and Turlach, B.A. (2000). On the LASSO and its dual. *Journal of Computational and Graphical Statistics*, 9, 319–337.

Polyak, B., Khlebnikov, M., and Shcherbakov, P. (2013). An lmi approach to structured sparse feedback design in linear control systems. In *European Control Conference (ECC)*, 833–838.

Tibshirani, R. (1996). Regression shrinkage and selection via the LASSO. *J. R. Statist. Soc. Ser. B*, 58(1), 267–288.

Zou, H. and Hastie, T. (2005). Regularization and variable selection via the Elastic Net. *Journal of the Royal Statistical Society, Series B*, 67, 301–320.